REMARKS

Rejection of Claims 1-19 Under 35 U.S.C. § 103(a)

Claims 1-19 are rejected under 35 U.S.C. § 103(a) as being obvious over Earhart, *et al.* in view of McGall, *et al.* The Examiner stated that the Declaration Under 37 C.F.R. § 1.132 is not sufficient to overcome the previous rejection of Claims 1-17 (Claims 18-19 were added by the prior Amendment), and raises a number of issues regarding the Declaration. Applicants will respond to each issue below.

The first issue raised by the Examiner relates to the methods by which the experiments described in the Declaration were conducted. The Examiner appears to be concerned that the date of the referenced publication is after the filing date of the instant application. As a practical matter, Applicants note that the experiments described in the Declaration were performed prior to the filing of the instant application and that the Declarant provided the McGall, *et al.* 2002 publication (previously submitted as "Exhibit A") as a convenient description of the methods used rather than reproducing the methods in their entirety. With respect to the claims, Claim 1 recites a method of oxidizing a phosphite ester linkage in a nucleic acid array to a phosphate linkage, comprising contacting the phosphite ester linkage with a solution of from about 0.005 M to about 0.05 M iodine in a mixture of water and organic solvent to form the phosphate linkage. It should be understood that the McGall, *et al.* 2002 publication provides a general method of nucleotide array preparation (see, for example, the paragraph bridging pages 24 and 25). The Declaration at page 3, second full paragraph states that:

> Synthesis of the arrays was conducted as described in "High-Density GeneChip Oligonucleotide Probe Arrays", Glenn H. McGall and Fred C. Christian, *Advances in Biochemical Engineering/Biotechnology* 77: 21-42 (2002), [previously] attached as "Exhibit A", ***with the exception that the iodine concentration in the oxidant solutions varied as shown below in Experiments 1 and 2.*** (emphasis added)

The Declaration clearly states that the oxidation step (referenced on page 25, line 2 of the McGall, *et al.* 2002 publication) has been modified to use a concentration of iodine within the claimed range. Thus, the exemplified method falls within the scope of Claim 1 because it uses a method of oxidizing a phosphite ester linkage in a nucleic acid array to a phosphate linkage, where the phosphite ester linkage is contacted with a concentration of iodine falling within the

claimed range. Accordingly, the Declaration should be given full consideration because the experiments described therein were conducted in accordance with the methods as claimed.

The Examiner also raises an issue with respect to the "average hybridization signal intensities" provided in the Declaration. The Examiner stated that there is no data regarding the range of hybridization signal intensities observed during the experiments. It should be understood that an "average" hybridization signal intensity does not refer to the mean signal intensity from a group of experiments. Instead, an average hybridization signal intensity represents the signal intensity of a single experiment that has been averaged over time for a particular region of a substrate. Therefore, there is no range of signal intensities to be reported.

The Declaration clearly shows that the signal intensities obtained with 0.10 M iodine are uniformly less than the signal intensities obtained with lower concentrations of iodine. In the discussion of unexpected results, it is not clear to Applicants what criteria the Examiner is using to judge the "unexpectedness" of the results. In particular, the Examiner has provided no explanation as to why the functional performance of nucleic acid arrays oxidized with 0.02 M iodine is considered to be unexpected while the functional performance of nucleic acid arrays oxidized with other concentrations is not regarded as unexpected.

The Declaration provides a clear standard for what constitutes an unexpected result. The Declarant states in Section 4 of the Declaration that it was previously thought that the functional performance of a nucleic acid array would not be compromised when the iodine concentration was sufficient to completely oxidize all phosphite esters to phosphate esters. This statement clearly assumes that the functional performance of a nucleic acid array ***reaches and maintains a maximum value as iodine concentration is increased***. Accordingly, an improvement in the functional performance of an array as iodine concentration is decreased from standard values (e.g., 0.10 M or greater) is unexpected. The nucleic acid arrays of Experiments 1 and 2 of the Declaration prepared using less than 0.10 M iodine all show at least a 20% improvement in the functional performance, as measured by the average hybridization signal. Because such improvements in functional performance are not predicted by conventional method of preparing nucleic acid arrays, the results shown in Experiments 1 and 2 are wholly unexpected.

The Examiner also raises an issue regarding use of the term "about" in the instant claims. This will be discussed below in conjunction with the Rejection Under 35 U.S.C. § 112, second paragraph.

Applicants have now responded to all issues the Examiner raised in considering the Declaration. Applicants have shown that the experiments described in the Declaration employ the claimed method. Applicants have also further clarified that the data provided in the Declaration are derived from two individual experiments (Experiments 1 and 2), and that an "average" hybridization signal intensity refers to an average taken over a discrete region of a chip from a single experiment. Applicants have also further clarified the standard by which unexpected results should be evaluated according to the Declaration. The Examiner has provided no evidence or reasoning to challenge this standard. The results described in the Declaration show at least a 20% improvement in the functional performance of the nucleic acid arrays prepared with the claimed concentration of iodine, and it is unexpected that improvements of this magnitude were obtained using a lower concentration of iodine than is conventionally used. Thus, Applicants maintain that the Declaration unambiguously demonstrates that the claimed method provides nucleic acid arrays with unexpectedly improved functional performance as compared with the cited art. Reconsideration and withdrawal of the rejection are respectfully requested.

Claims 4 and 15-17

The Examiner indicated in the telephonic interview of December 3, 2002 that Claims 4 and 15-17 should have been indicated as allowable, based solely upon the data provided in the specification. The Examiner, at page 3 of the Office Action, reiterates that the McGall Declaration provides unexpectedly higher hybridization intensity values for the 0.02 M concentration. However, Applicants note that these claims continue to stand rejected; clarification is requested.

Rejection of Claims 18-19 Under 35 U.S.C. § 112, First Paragraph

Claims 18-19 are rejected under 35 U.S.C. § 112, first paragraph, for allegedly failing to comply with the written description requirement. The Examiner stated that these claims contain

subject matter which was not described in the specification. In particular, the Examiner stated that there is no support for the ranges of "from about 0.01 M to about 0.05 M iodine" and "from about 0.02 M to about 0.05 M iodine". Applicants respectfully traverse this rejection.

The MPEP provides a discussion of the written description requirement as it applies to claimed ranges and subranges. MPEP § 2163.05 states:

> With respect to changing numerical range limitations, the analysis must take into account which ranges one skilled in the art would consider inherently supported by the discussion in the original disclosure. In the decision in *In re Wertheim*, 541 F.2d 257, 191 USPQ 90 (CCPA 1976), the ranges described in the original specification included a range of "25%- 60%" and specific examples of "36%" and "50%." A corresponding new claim limitation to "at least 35%" did not meet the description requirement because the phrase "at least" had no upper limit and caused the claim to read literally on embodiments outside the "25% to 60%" range, however a limitation to "between 35% and 60%" did meet the description requirement.

The instant application discloses the range of from about 0.005 M to about 0.05 M in numerous locations, for example, at page 19, lines 26-31. The instant specification also discloses preferred embodiments where about 0.02 M iodine is present, for example, at page 21, lines 17-25. Based on the discussion in MPEP § 2163.05, a subrange that falls within a disclosed range is considered to be inherently disclosed by the specification. Thus, the ranges of Claims 18 and 19 (about 0.01 M to about 0.05 M and about 0.02 M to about 0.05 M, respectively) clearly comply with the written description requirement. Reconsideration and withdrawal of the rejection are respectfully requested.

Rejection of Claims 1-19 Under 35 U.S.C. § 112, Second Paragraph

Claims 1-19 are rejected under 35 U.S.C. § 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention. The Examiner stated that one of ordinary skill in the art would not be able to ascertain the full scope of the claimed invention without the definition of the term "about".

Applicants respectfully disagree that the term "about" renders the instant claims indefinite. MPEP § 2173.05(b) provides the following discussion of the use of the term "about":

> The term "about" used to define the area of the lower end of a mold as between 25 to about 45% of the mold entrance was held to be clear, but flexible. *Ex parte Eastwood*, 163 USPQ 316 (Bd. App. 1968). Similarly, in *W.L. Gore & Associates, Inc. v. Garlock, Inc.*, 721 F.2d 1540, 220 USPQ 303 (Fed. Cir. 1983), the court held that a limitation defining the stretch rate of a plastic as "exceeding about 10% per second" is definite because infringement could clearly be assessed through the use of a stopwatch. However, the court held that claims reciting "at least about" were invalid for indefiniteness where there was close prior art and there was nothing in the specification, prosecution history, or the prior art to provide any indication as to what range of specific activity is covered by the term "about." *Amgen, Inc. v. Chugai Pharmaceutical Co.*, 927 F.2d 1200, 18 USPQ2d 1016 (Fed. Cir. 1991).

The instant claims present a situation analogous to *ex parte Eastwood* and *W.L. Gore & Associates, Inc. v. Garlock, Inc.* Like *W.L. Gore & Associates*, infringement can be readily assessed by measuring the concentration of iodine in solution (e.g., via spectrophotometric methods). Unlike *Amgen, Inc.*, the closest prior art of record discloses a concentration of iodine that is approximately 2 times greater than the upper bound of the claimed range. It is respectfully submitted that a 2-fold difference does not constitute "close prior art". Moreover, the use of "about" in the claims recognizes that there is some experimental error in preparing solutions having a desired concentration and measuring the concentration of a solution. It would be inequitable to require Applicants to delete "about" from the claims.

Applicants maintain that the claimed ranges clearly delineate the bounds of the invention and that a potential infringer could readily whether a particular method was encompassed by the instant claims (i.e., by determining the iodine concentration). Reconsideration and withdrawal of the rejection are respectfully requested.

Information Disclosure Statement

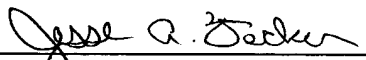
An Information Disclosure Statement (IDS) was filed on May 25, 2001 and a Supplemental IDS was filed on March 17, 2003. To date, Applicants have not received a copy of the Form PTO-1449 initialed by the Examiner to indicate consideration of the cited references. Applicants again respectfully request that the Examiner enter the IDS and Supplemental IDS in the record and return a copy of the initialed Form PTO-1449 with the next communication.

CONCLUSION

In view of the above amendments and remarks, it is believed that all claims are in condition for allowance, and it is respectfully requested that the application be passed to issue. If the Examiner feels that a telephone conference would expedite prosecution of this case, the Examiner is invited to call the undersigned.

Respectfully submitted,

HAMILTON, BROOK, SMITH & REYNOLDS, P.C.

By 

Jesse A. Fecker
Registration No. 52,883
Telephone: (978) 341-0036
Facsimile: (978) 341-0136

Concord, MA 01742-9133
Dated: 8-28-03

High-Density GeneChip Oligonucleotide Probe Arrays

Glenn H. McGall[1] · Fred C. Christians[2]

[1] Affymetrix, Inc., 3380 Central Expressway, Santa Clara, CA 95051, USA.
E-mail: glenn_mcgall@affymetrix.com
[2] Affymetrix, Inc., 3380 Central Expressway, Santa Clara, CA 95051, USA.
E-mail: fred_christians@affymetrix.com

High-density DNA probe arrays provide a highly parallel approach to nucleic acid sequence analysis that is transforming gene-based biomedical research. Photolithographic DNA synthesis has enabled the large-scale production of GeneChip probe arrays containing hundreds of thousands of oligonucleotide sequences on a glass "chip" about 1.5 cm^2 in size. The manufacturing process integrates solid-phase photochemical oligonucleotide synthesis with lithographic techniques similar to those used in the microelectronics industry. Due to their very high information content, GeneChip probe arrays are finding widespread use in the hybridization-based detection and analysis of mutations and polymorphisms ("genotyping"), and in a wide range of gene expression studies.

Keywords: GeneChip array, Oligonucleotide probe array, Photolithography, Gene expression monitoring, Genotyping

Abbreviations

CE	2-cyanoethyl
DMT	4,4′ dimethoxytriphenylmethyl
DLP	digital light processor

EXHIBIT
A

Advances in Biochemical Engineering/
Biotechnology, Vol. 77
Managing Editor: T. Scheper

DTT	dithiothreitol
HPLC	high performance liquid chromatography
kb	kilobase
MeNPOC	α-methyl-6-nitropiperonyloxycarbonyl
NNEOC	1-(8-nitronaphth-1-yl)ethyloxycarbonyl
NPPOC	2-(2-nitrophenyl)propyloxycarbonyl
NTP	nucleoside 5′-triphosphate
PAG	photo-acid generator
PEO	poly(ethylene oxide)
PIV	pivaloate
pTs	para-toluenesulfonate
PYMOC	1-pyrenylmethyloxycarbonyl
SNP	single-nucleotide polymorphism
TCEP	tris(2-carboxyethyl)phosphine

1 Introduction

High-density polynucleotide probe arrays are among the most powerful and versatile tools for accessing the rapidly growing body of sequence information that is being generated by numerous public and private sequencing efforts. Consequently, this technology is having a major impact on biological and biomedical research [1, 2]. These arrays are essentially large sets of nucleic acid probe sequences immobilized in defined, addressable locations on the surface of a substrate capable of accessing large amounts of genetic information from biological samples in a single hybridization assay. In a typical application [2], DNA or RNA "target" sequences of interest are isolated from a biological sample using standard molecular biology protocols. The sequences are fragmented and labeled with fluorescent molecules for detection, and the mixture of labeled sequences is applied to the array, under controlled conditions, for hybridization with the surface probes. The array is then imaged with a fluorescence-based reader to locate and quantify the binding of target sequences from the sample to complementary sequences on the array, and software reconstructs the sequence data and presents it in a format determined by the application. Thus, in addition to the arrays themselves, the Affymetrix GeneChip system provides a fluidics station for performing reproducible, automated hybridization and wash functions; a high-resolution scanner for reading the fluorescent hybridization image on the arrays; and software for processing and querying the data (Fig. 1).

GeneChip technology is distinguishable from other microarray methods in that oligonucleotide probe sequences are photolithographically synthesized, in a parallel fashion, directly on a glass substrate. In a minimum number of synthetic steps, arrays containing hundreds of thousands of different probe sequences, typically 20 or 25 bases in length, can be generated at densities in the order of 10^5–10^6 sequences/cm^2 (Fig. 2). This capability, deliverable on a commercial production scale, is well beyond that of any other technology currently

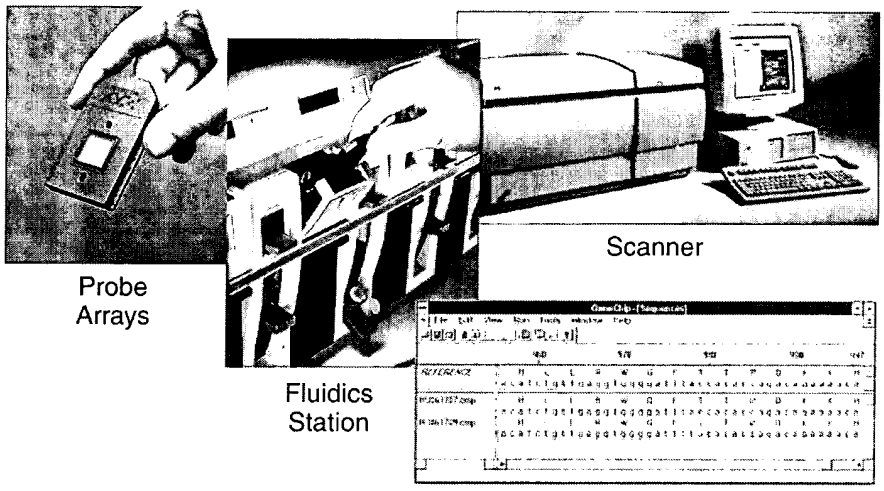


Fig. 1. GeneChip system overview

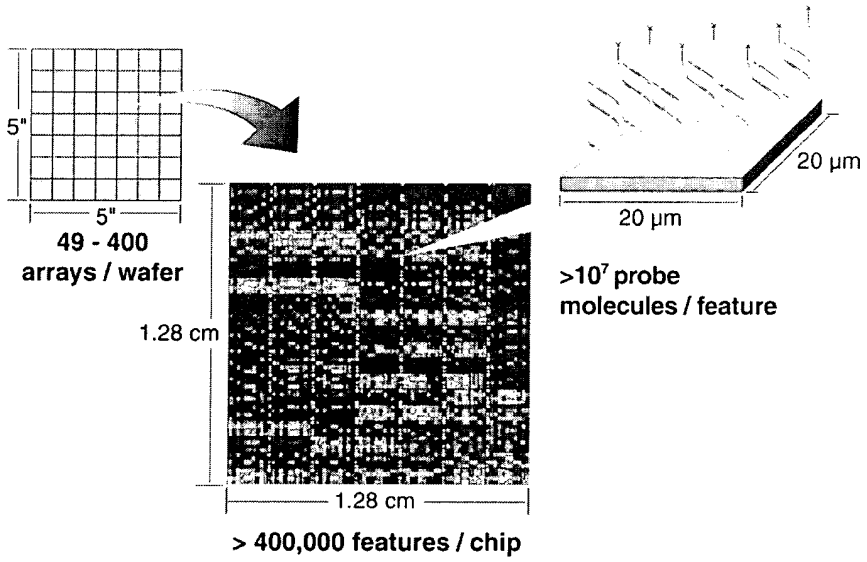


Fig. 2. Wafer-scale GeneChip production specifications

available and allows unprecedented amounts of sequence information to be encoded in the arrays.

Other array construction technologies, such as micropipetting or inkjet printing, rely on mechanical devices to deliver minute quantities of reagents to predefined regions of a substrate in a sequential fashion. In contrast, the photolithographic synthesis process is highly parallel in nature, making it intrinsically robust and scalable. This provides significant flexibility and cost advantages in terms of materials management, manufacturing throughput, and quality control. To researchers the benefits are a high degree of reliability, uniformity of array performance, and an affordable price.

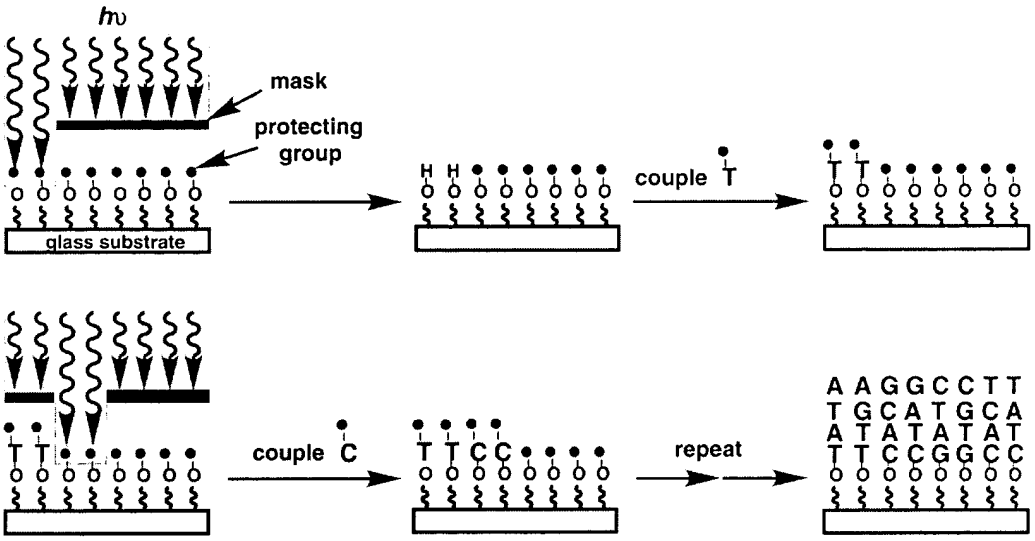


Fig. 3. Photolithographic synthesis of oligonucleotide arrays

2 Array Production Technology

The advent of DNA array technology has relied on the development of methods for fabricating arrays with sufficiently high information content and density in a rapid, reproducible and economic fashion. Light-directed synthesis [3–7] has made it possible to manufacture arrays containing hundreds of thousands of oligonucleotide probe sequences on glass "chips" little more than 1 cm^2 in size, and to do so on a commercial scale. In this process, 5′- or 3′-terminal protecting groups are selectively removed from growing oligonucleotide chains in predefined regions of a glass support, by controlled exposure to light through photolithographic masks (Fig. 3).

2.1 Substrate Preparation and General Approach

Prior to photolithographic synthesis, planar glass substrates are covalently modified with a silane reagent to provide a uniform layer of covalently bonded hydroxyalkyl groups on which oligonucleotide synthesis can be initiated (Fig. 4). In a second step, a photo-imagable layer is added by extending these synthesis sites with a poly(ethylene oxide) linker which has a terminal photolabile hydroxyl protecting group. When specific regions of the surface are exposed to light, synthesis sites within these regions are selectively deprotected, and thereby "activated" for the addition of nucleoside phosphoramidite building blocks.

These nucleotide precursors, also protected at the 5′ or 3′ position with a photolabile protecting group, are applied to the substrate, where they react with the surface hydroxyl groups in the pre-irradiated regions. The monomer coupling step is carried out in the presence of a suitable activator, such as tetrazole or di-

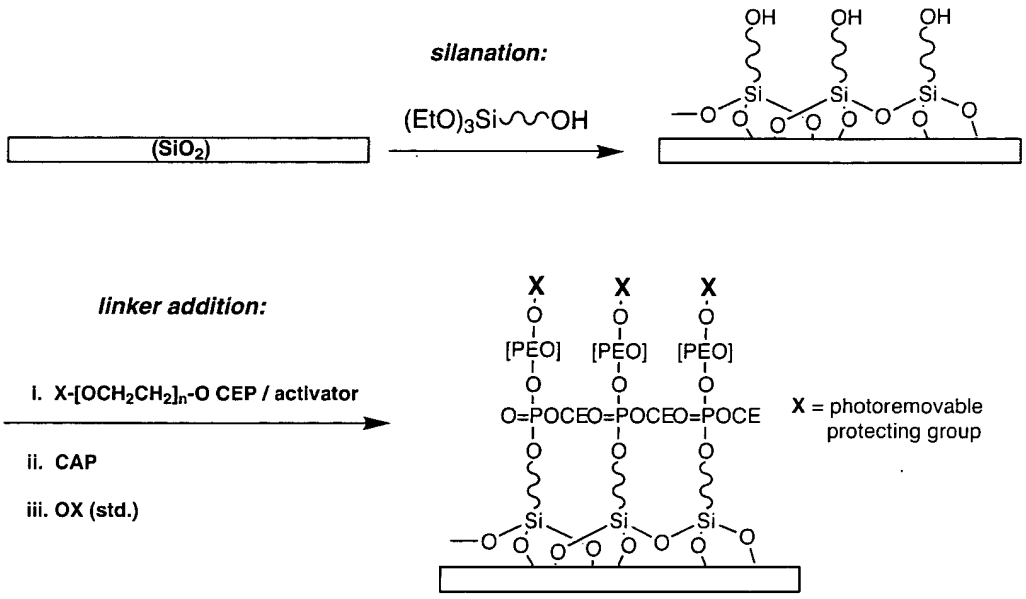


Fig. 4. Chemical preparation of glass substrates for light-directed synthesis of oligonucleotide arrays

cyanoimidazole. The coupling reaction is followed by conventional capping and oxidation steps, which also use standard reagents and protocols for oligonucleotide synthesis [5, 7]. Alternating cycles of photolithographic deprotection and nucleotide addition are repeated to build the desired two-dimensional array of sequences, as described in Fig. 3.

Semiautomated cleanroom manufacturing techniques, similar to those used in the microelectronics industry, have been adapted for the large-scale commercial production of GeneChip arrays in a multi-chip wafer format (Fig. 5). Each wafer contains 49–400 replicate arrays, depending on the size of the array,

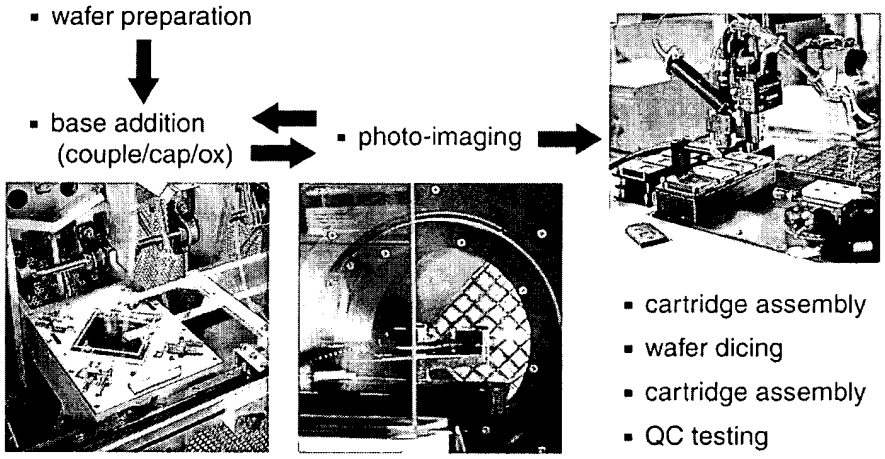


Fig. 5. Automated array manufacturing

and multiple-wafer lots can be processed together in a procedure that takes less than 24 h to complete. Multiple lots are processed simultaneously on independent production synthesizers operating continuously. After a final chemical deprotection, finished wafers are diced into individual arrays, which are finally mounted in injection-molded plastic cartridges for single-use application (see Fig. 1).

2.2 Photolithography

The photolithographic process provides a very efficient route to high-density arrays by allowing parallel synthesis of large sets of probe sequences. The number of required synthetic steps to fabricate an array is dependent only on the length of the probes, not the number of probes. A complete set, or any subset, of probe sequences of length "n" requires $4 \times n$ synthetic steps, at most. Masks can be designed to make arrays of oligonucleotide probe sequences for a variety of applications. Most arrays are comprised of custom-designed sets of probes 20–25 bases in length, and optimized masking strategies allow such arrays to be completed in as few as $3n$ steps.

The spatial resolution of the photolithographic process determines the maximum achievable density of the array and therefore the amount of sequence information that can be encoded on a chip of a given physical dimension. A contact lithography process (Fig. 3) is used to fabricate GeneChip arrays with individual probe features that are 18×18 microns in size. Between 49 and 400 identical arrays are produced simultaneously on $5'' \times 5''$ wafers. For the largest-format chip currently in production (1.6 cm^2), this provides wafers of 49 individual arrays containing more than 400,000 different probe sequences each. For arrays containing fewer probe sequences, this feature size enables more replicate arrays to be fabricated on each wafer. The technology has proven capability for fabricating arrays with densities greater than 10^6 sequences/cm^2, corresponding to features less than 10×10 microns in size. This level of miniaturization is beyond the current reach of other technologies for array fabrication.

2.3 Light-Directed Synthesis Chemistry

The current manufacturing process employs nucleoside monomers protected with a photoremovable 5′-(α-methyl-6-nitropiperonyloxycarbonyl), or "MeNPOC" group [4, 5], depicted in Fig. 6, which offers a number of advantages for large-scale manufacturing. These phosphoramidite monomers are relatively inexpensive to prepare, and photolytic deprotection is induced by irradiation at near-UV wavelengths ($\phi \sim 0.05$; $\lambda_{max} \sim 350$ nm) so that photochemical modification of the oligonucleotides, which absorb energy at lower wavelengths, can be avoided. The photolysis reaction involves an intramolecular redox reaction and does not require any special solvents, catalysts or coreactants. Since the photolysis can be performed "dry", high-contrast contact lithography can be used to achieve very high-resolution imaging. Complete photodeprotection requires less

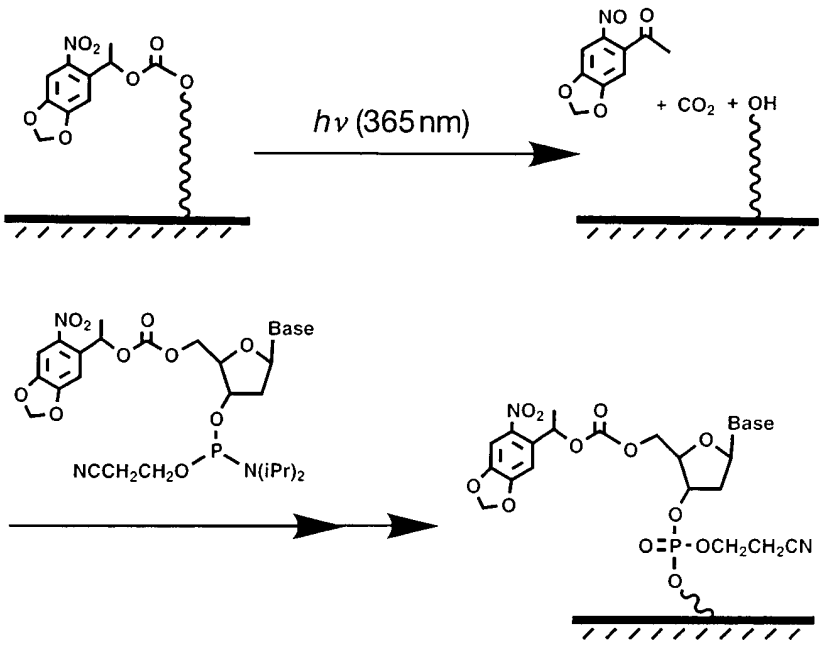


Fig. 6. Light-directed oligonucleotide synthesis cycle using MeNPOC photolabile phosphoramidite building blocks

than 1 min using filtered I-line (365 nm) emission from a commercial collimated mercury light source.

Photochemical deprotection rates and yields for oligonucleotide synthesis can both be monitored directly on planar supports using procedures based on surface fluorescence [5]. We have also developed a sensitive assay in which test sequences are synthesized on a support designed to allow the cleavage and direct quantitative analysis of labeled oligonucleotide products using ion-exchange high performance liquid chromatography (HPLC) with fluorescence detection [8, 9]. This method involves photolithographic synthesis of test sequences after the addition of a base-stable disulfide linker and a fluorescein monomer to the support (Fig. 7). The disulfide linker remains intact throughout the synthesis and deprotection, but can be subsequently cleaved under reducing conditions to release the synthesis products, all of which are uniformly labeled with a 3′-fluorescein tag. The labeled oligonucleotide synthesis products are then analyzed using HPLC or capillary electrophoresis with fluorescence detection, enabling direct quantitative analysis of synthesis efficiency. The sensitivity of fluorescence is a key feature of this methodology, since the quantities of DNA synthesis products on flat substrates are very low (1–100 pmol/cm^2 [9]), and difficult to analyze by other means.

The average stepwise efficiency of the light-directed oligonucleotide synthesis process is limited by the yield of the photochemical deprotection step which, in the case of MeNPOC nucleotides, is 90%–94% [5]. The other chemical reactions involved in the base addition cycles (coupling, capping, oxidation) use reagents in a vast excess over surface synthesis sites, and, provided that sufficient reagent concentrations and time are allowed for completion, they are es-

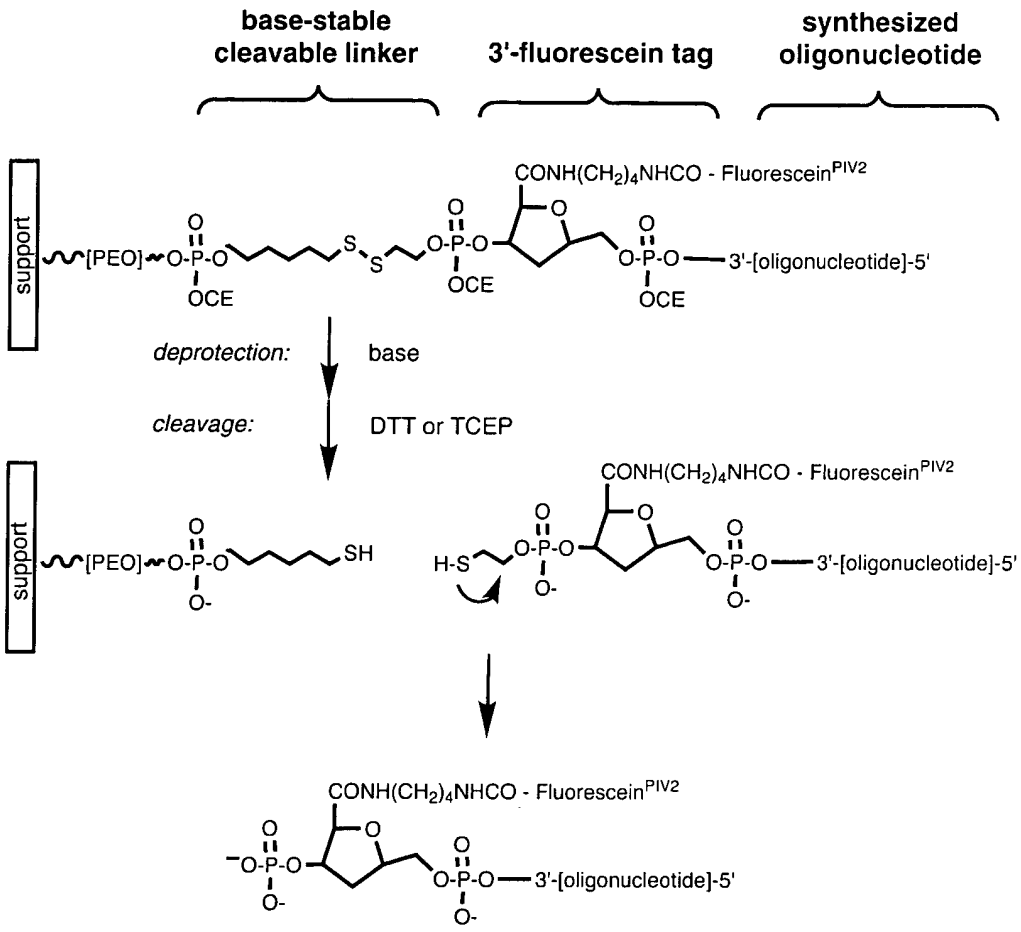


Fig. 7. Method for fluorescent labeling and cleavage of photolithographically synthesized oligonucleotides allows quantitative analysis by HPLC

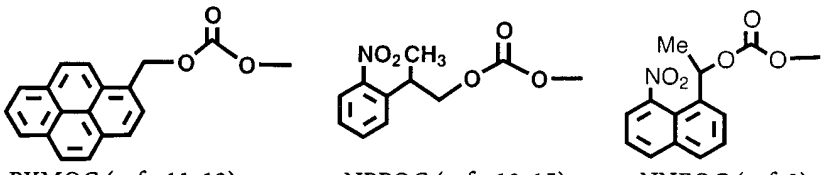
sentially quantitative. However, the sub-quantitative photolysis yields lead to incomplete or "truncated" probes, with the desired full-length sequences representing, in the case of 20mer probes, approximately 10% of the total synthesis products.

For a number of reasons, the presence of truncated probe impurities has a relatively minor impact on the performance characteristics of arrays when they are used for hybridization-based sequence analysis. Firstly, the silanating agents used in this process provide an abundance of initial surface synthesis sites (>100 pmol/cm^2), so that the final "concentration" of completed probes on the support remains high. Thus, each of the 20×20 micron features on a typical array contains over 10^7 full-length oligonucleotide molecules (Fig. 2). It should be noted that there is an optimum surface probe density for maximum hybridization signal and discrimination. Thus, an increase in the synthetic yield through alternative chemistries or processes, while increasing the surface concentration of full-length probes, can actually reduce the quality of the hybridization data [10]. This is due to steric and electrostatic repulsive effects that result when oligonucleotide molecules are too densely spaced on the support. Secondly, ar-

ray hybridizations are typically carried out under stringent conditions so that hybridization to significantly shorter (< approximately $n-4$) oligomers is negligible. Longer truncated sequences are of low abundance, and contribute little to the total hybridization signal in a probe feature. Furthermore, the truncated probes retain correct sequences, and any residual binding will be to the target sequences for which they were designed, albeit with slightly lower specificity. These factors, combined with the use of comparative intensity algorithms for data analysis [6], allow highly accurate sequence information to be "read" from these arrays with single-base resolution.

2.4 Future Enhancements

Improvements in Synthetic Yield. Several alternative photolabile protecting group chemistries have been described which may also be applicable to light-directed DNA array synthesis [9–15]. Some are capable of providing stepwise coupling yields in excess of 96%, and several examples are shown in Fig. 8. Achieving high synthetic yields with these alternative chemistries typically requires that a layer of solvent or catalyst be maintained over the substrate during the photodeprotection step. However, this has the drawback of adding significant complexity and cost to the manufacturing process. Furthermore, when solvent is required, the light image must be projected through the substrate from the reverse side, and image quality is somewhat degraded, thus limiting the achievable density of arrays that can be made this way. Nonetheless, certain array applications that have more stringent probe purity requirements could benefit substantially from improvements in photochemical synthesis yield. Such applications would include those in which, after hybridization, the probe-target duplexes are used as a platform for further reactions or analyses. For example, methods based on template-directed enzymatic probe extension, ligation or cleavage have been suggested as a means of improving allelic discrimination in hybridization-based mutation and polymorphism detection on arrays [16]. For this reason, we are currently developing a new generation of reagents for photolithographic synthesis that provide high synthetic yields without impacting the cost or lithographic performance of the current manufacturing process. Some of these biochemical assay formats will also require probe array synthesis to proceed in the 5′–3′ direction so that the probes will be attached to the support at the 5′-terminus. This can be achieved through the use of 3′-photo-activatable 5′-phosphoramidite building blocks [7].



PYMOC (refs. 11, 12) NPPOC (refs. 13, 15) NNEOC (ref. 9)

Fig. 8. Alternate photoremovable protecting groups for photolithographic oligonucleotide synthesis

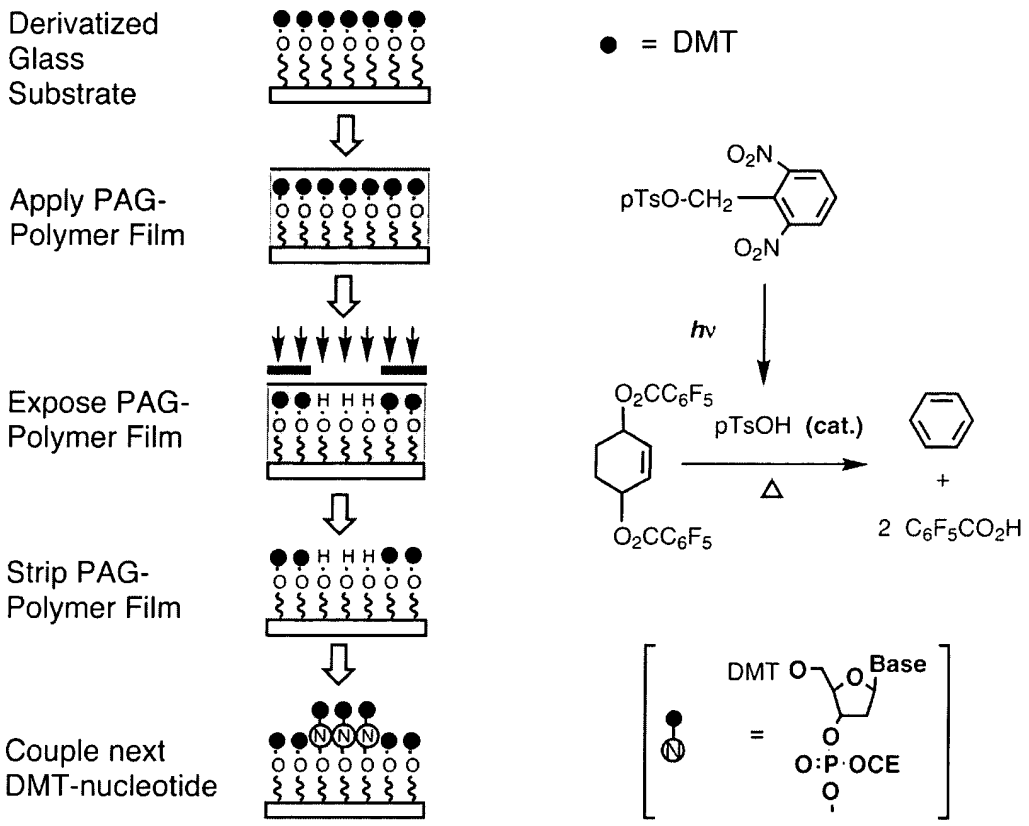


Fig 9. DNA probe array synthesis using photoacid generation in a polymer film to remove acid-labile DMT protecting groups

Improvements in Photolithographic Resolution. In order to achieve higher spatial resolution, as well as synthetic yields and photolysis rates, we have developed novel photolithographic methods for fabricating DNA arrays which exploit polymeric photo-resist films as the photoimageable component [17–20]. One of the advantages of the photo-resist approach is that it can utilize conventional 4,4′-dimethoxytrityl (DMT)-protected nucleotide monomers. These processes can also make use of chemical amplification of photo-generated catalysts to achieve higher contrast and sensitivity (shorter exposure times) than conventional photo-removable protecting groups. In this process, a polymeric thin film, containing a chemically amplified photo-acid generating (PAG) system, is applied to the glass substrate. Exposure of the film to light creates localized development of an acid catalyst in the film adjacent to the substrate surface, resulting in direct removal of DMT protecting groups from the oligonucleotide chains (Fig. 9). This process has provided stepwise synthetic yields >98%, photolysis speeds at least an order of magnitude faster than that achievable with photo-removable protecting groups, and photolithographic resolution capability well below 10 microns. This technology will enable the production of arrays with much higher information content than is currently attainable.

Flexible Custom Array Fabrication. Another recent development has been the application of programmable digital micromirror devices, or "digital light processors" (DLPs), for photolithographic imaging, which offers a flexible approach to custom photolithographic array fabrication [21]. These devices were originally developed for digital image projection in consumer electronics products. They are essentially high-density arrays of switchable mirrors that reflect light from a source into an optical system that focuses and projects the reflected image. By using DLPs for photolithographic array synthesis, custom designs can be fabricated in a relatively short time, without the need for custom chrome-glass mask sets. This approach may offer an advantage to researchers who wish to vary designs frequently, and use only a small number of arrays of a given design. It should be noted that the standard lithographic approach using chrome-on-glass masks, which is ideal for mass production of standardized arrays, is also being adapted to the cost-effective production of smaller quantities of variable-content arrays. This is achieved through the use of high-throughput mask design and fabrication capabilities, combined with new strategies that dramatically reduce the number of masks required to synthesize custom arrays.

3 Applications

GeneChip oligonucleotide probe arrays are used to access genetic information contained in both the RNA (gene expression monitoring) and DNA (genotyping) content of biological samples. Many different GeneChip products are now available for gene expression monitoring and genotyping complex samples from a variety of organisms. The ability to simultaneously evaluate tens of thousands of different mRNA transcripts or DNA loci is transforming the nature of basic and applied research, and the range of application of DNA probe arrays is expanding at an accelerating pace. Current information on Affymetrix products and specifications is available at the website (www.affymetrix.com/products). A number of representative applications of these arrays are discussed below.

3.1 Gene Expression Monitoring

Currently, the most popular application for oligonucleotide microarrays is in monitoring cellular gene expression. Standard GeneChip arrays are encoded with public sequence information, but custom arrays are also designed from proprietary sequences. Figure 10 depicts how a gene expression array interrogates each transcript at multiple positions. This feature provides more accurate and reliable quantitative information relative to arrays that use a single probe, such as a cDNA, PCR amplicon, or synthetic oligonucleotide for each transcript. Two probes are used at each targeted position of the transcript, one complementary (perfect match probe), and one with a single base mismatch at the central position (mismatch probe). The mismatch probe is used to estimate and correct for both background and signal due to non-specific hybridization. The number of transcripts evaluated per probe array depends on chip size, the individual probe

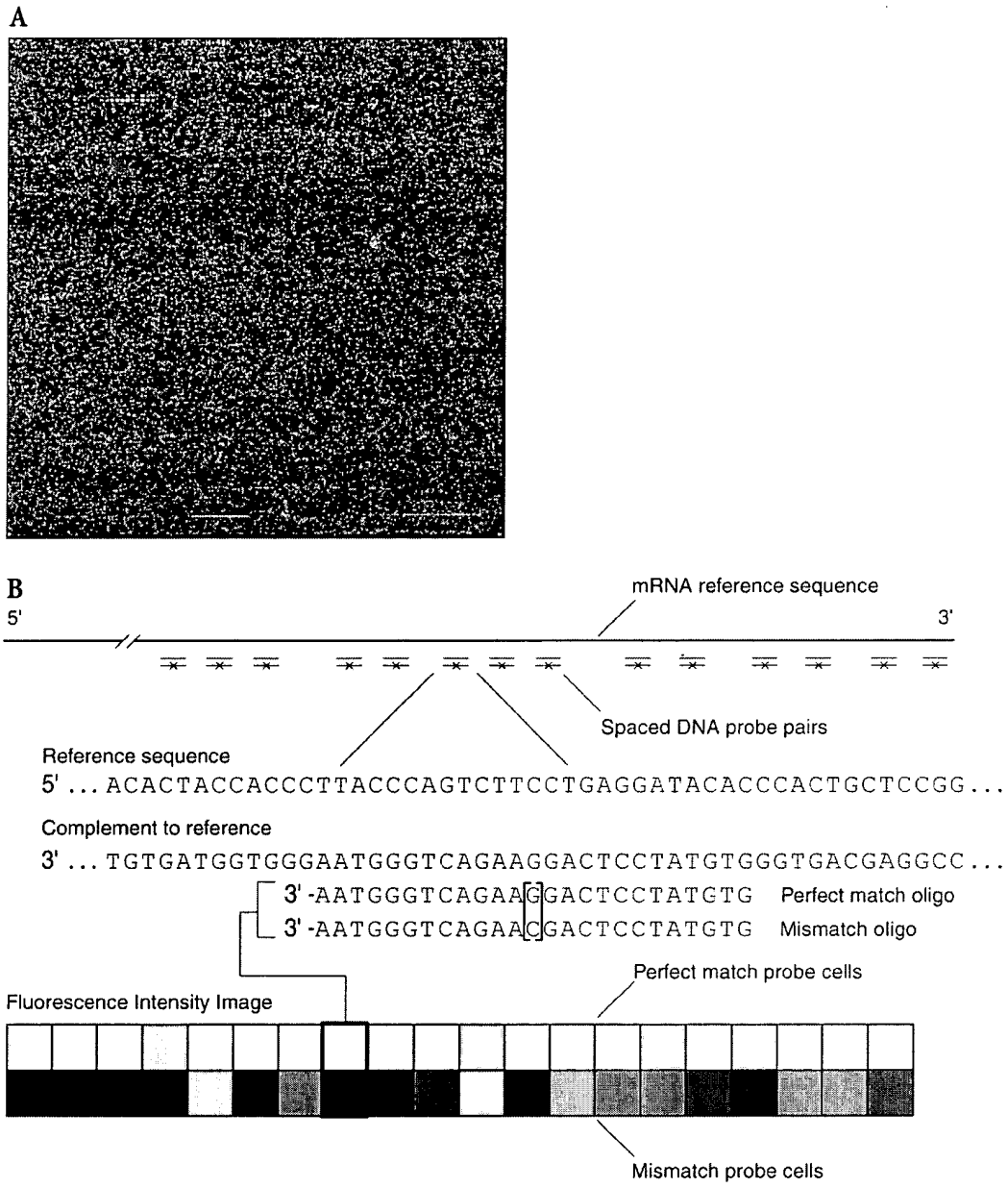


Fig. 10. Gene expression monitoring with oligonucleotide arrays. **A** An image of a hybridized 1.28 × 1.28 cm HuGeneFL array, with 20 probe pairs for each of approximately 5000 full-length human genes. **B** Probe design. To control for background and cross-hybridization, each perfect match probe is partnered with a probe of the same sequence except containing a central mismatch. Probes are usually 25mers, and are generally chosen to interrogate the 3′ regions of eukaryotic transcripts to mitigate the consequences of partially degraded mRNA. (Reprinted with permission from [6])

feature size, and the number of probes dedicated to each transcript. A standard 1.28 × 1.28 cm probe array, with individual 18 × 18 micron features, and 16 probe pairs per probe set, can interrogate over 20,000 transcripts. This number is steadily increasing as manufacturing improvements shrink the feature size, and

as improved sequence information and probe selection rules allow reductions in the number of probes needed for each transcript.

Arrays are now available to examine entire transcriptomes from a variety of organisms including several bacteria, yeast, drosophila, arabidopsis, mouse, rat, and human. Instead of monitoring the expression of a small subset of selected genes, researchers can now monitor the expression of all or nearly all of the genes for these organisms simultaneously, including a large number of genes of unknown function. Numerous facets of biology and medicine are being explored using this powerful new capability. Gene function has been explored by studying changes in expression levels throughout the cell cycle [22, 23]. Genetic pathways can be examined in great detail by monitoring the downstream transcriptional effects of inducing specific genes in cell culture, and the effects of drug treatment on gene expression levels can be surveyed [24]. Expression arrays have also been used to screen thousands of genes to identify markers for human diseases such as cancer [25], muscular dystrophy [26], diabetes [27], or for aging [28, 29].

One important area of research that is benefiting greatly from GeneChip technology is cancer profiling, wherein gene expression monitoring is used to classify tumors in terms of their pathologies and responses to therapy. Understanding the variation among cancers is the key to improving their treatment. For example, a prostate tumor may be essentially harmless for 20 years in one patient, while an apparently similar tumor in another patient can be fatal within several months. One patient's lymphoma may respond well to chemotherapy while another will not. This variation of pathologies has motivated oncologists to assemble an impressive body of information to help classify tumors based on numerous histological, molecular, and clinical parameters. This has required a massive effort by thousands of highly skilled and dedicated scientists over the past few decades. It is clear, however, that there is a need for improvement in tumor classification in terms of both understanding variations among tumors, and assigning tumors to appropriate classes.

Two recent studies demonstrate the utility of GeneChip technology to cancer classification. In the first report [30], expression levels were measured in colon adenocarcinomas and normal colon tissues, and patterns differentiating the two sample types were revealed by two-way clustering analysis. It was found that ribosomal proteins were expressed at consistently higher levels in tumors than in normal tissue, thereby providing a small set of characteristic markers with large expression differences. Muscle-specific genes provided another distinguishing set of markers. These genes were expressed at a higher level in normal tissue, probably because of the tissue composition of the samples, largely epithelial tissue for the tumors and mixed tissue for the normal. In addition to identifying these two small sets of genes with relatively large expression differences, the authors report that examination of large sets of genes with even small differences in expression could reliably classify a sample as tumor or normal. Although the confidence level of an individual gene might be low, extensive expression analysis using thousands of genes reveals subtle, systematic differences with high confidence. Such an expression profile database constructed from samples of known types would be useful for class prediction, that is, classifying additional unknown samples.

In a second report [31], class predictions were tested directly using a database built from two different types of leukemia. The current clinical tests used to distinguish ALL (acute lymphoblastic leukemia) from AML (acute myeloid leukemia), while useful, are imperfect and painstaking. Proper identification of these cancer types is critical because the treatments they require are quite different. In the first set of experiments, the authors examined 27 ALL and 11 AML samples to create a gene expression profile database. Statistical analysis identified about 1100 genes (of the 6817 examined) that correlated well with the ALL-AML class distinction. Many of the most highly correlated genes had been previously implicated in cancer. The 50 most informative genes were used to accurately cross-validate 36 of 38 of the original samples (the other 2 calls were uncertain), and were also good predictors for an independent set of leukemia samples (29 of 34 were called correctly; the other 5 were uncertain). The test on independent samples is especially convincing given the more varied nature of the samples: two different RNA isolation methods were used by different laboratories; some samples were isolated from peripheral blood, some from bone marrow; some AML patients were adults, and some were children. Thus, the panel of the 50 most informative genes served as a strong class prediction set, and the authors reported similar results using 10–200 genes. The authors point out, however, that for other sample groups the number of genes needed for accurate class prediction may vary, and that no single gene correlated completely with class type.

Golub et al. [31] went on to address the issue of class discovery, that is, the identification of new cancer classes. The statistical tool called self-organizing maps was employed to determine whether the original data set could be subdivided beyond the ALL-AML categories. The AML samples again clustered together, but the ALL samples were now split into two groups that were subsequently shown by immunotyping to be of B-cell or T-cell origin. Although this ALL subdivision was previously known, the clustering analysis would have discovered it even if it had not been known.

Expression profiling coupled with appropriate statistical analysis holds promise not only in cancer classification, but also by extension to many other areas of disease research and management. Transcription profile databases may be assembled from samples that differ by tissue source, disease state or progression, morbidity/mortality, response to drugs and other treatments, and countless other variables. New patterns may be revealed and disease classes refined or discovered. Patients may be more finely stratified in clinical trials so that the success of treatments can be better judged, and the expectation is that the diagnosis and treatment of disease will improve substantially as a result.

3.2 Genotypic Analysis

As the human genome project finishes the first complete blueprint of the human genome, there is tremendous interest in identifying the variations in DNA sequences between individuals and relating these variations to phenotypes. It is of particular interest to understand how subtle sequence differences are associated

with disease. Oligonucleotide arrays are well suited to probe these variations, particularly single-nucleotide substitutions and, to a lesser degree, short deletions and insertions.

Oligonucleotide arrays are currently used primarily for two types of genotyping analysis. *Arrays for mutation or variant detection* (Fig. 11) are used to screen sets of contiguous sequence for single-nucleotide differences. Given a reference sequence, the basic design of genotyping arrays is quite simple: four probes, varying only in the central position and each containing the reference sequence at all other positions, are made to interrogate each nucleotide of the reference sequence. The target sequence hybridizes most strongly to its perfect complement on the array, which in most cases will be the probe corresponding to the reference sequence, but, in the case of a nucleotide substitution, this will be one of the other three probes. Up to 50 kb of sequence can thus be determined with 200,000 probes, or 400,000 probes if both DNA strands are interrogated. Impending decreases in array feature size (see Sect. 2.4) will extend this capacity further. Mutation detection arrays have been used to analyze the entire 16.5 kb sequence of mitochondrial DNA samples [32], the 9.2 kb coding sequence of the ATM gene [33], BRCA1 coding mutations [34, 35], p53 mutations [36, 37], cytochrome P450 variants involved in drug metabolism [38], and HIV sequence variants [39, 40], among others. The last three arrays are commercial products currently available from Affymetrix.

The other main type of genotyping performed with oligonucleotide arrays is *SNP analysis*, that is, the genotyping of biallelic single-nucleotide polymor-

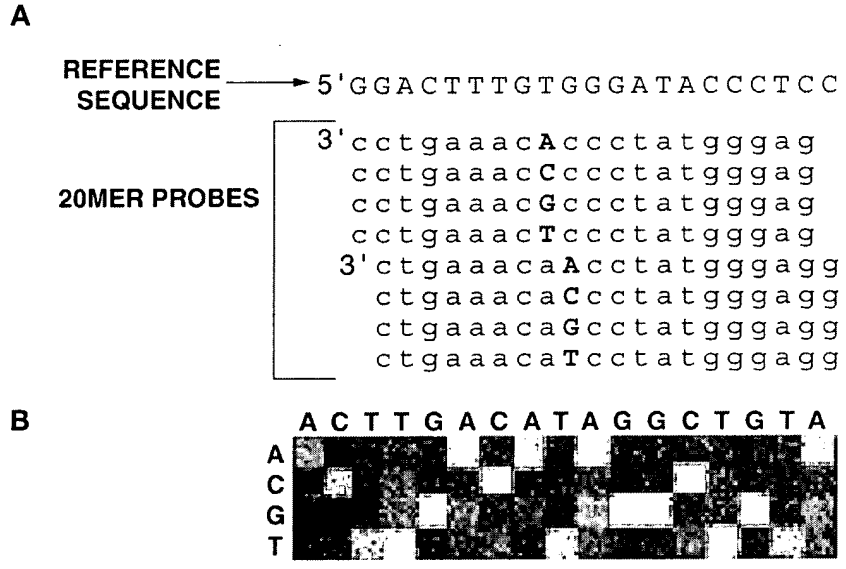


Fig. 11. Resequencing array for sequence variation detection. **A** Each base of a given reference sequence is represented by four probes, usually 20mers, that are identical to each other with the exception of a single centrally located substitution (*bold*). Shown are probe sets targeted to two adjacent positions of the reference sequence. **B** The target sequence is determined by hybridization intensities, with the probe complementary to the target providing the strongest signal. (Reprinted with permission from Warrington JA et al (2000) In: Microarray biochip technology. Biotechniques Books, p 122)

phisms. Because SNPs are the most common source of variation between individuals, they serve not only as landmarks to create dense genome maps, but also as markers for linkage and loss of heterozygosity studies. Large numbers of publicly available SNPs – over one million to date – have been found using dideoxy sequencing as well as mutation detection arrays [41–43].

In addition to mutation detection arrays, at least two other types of oligonucleotide arrays can be used for SNP analysis. The "HuSNP" assay allows nearly 1500 SNP-containing regions of the human genome to be amplified in just 24 multiplex PCRs and then hybridized to a single HuSNP array. The SNPs cover all 22 autosomes and the X chromosome. Many of the 1500 SNPs were discovered using mutation detection arrays [41]. The probe strategy for a SNP array is shown in Fig. 12. The probes for each SNP on the HuSNP array interrogate not only the two alleles of the SNP position, but also three or four positions flanking the SNP; the redundant data are of higher quality for the same reasons that the use of multiple probes improves gene expression monitoring array data. SNP arrays have also proven useful for loss of heterozygosity studies [44, 45].

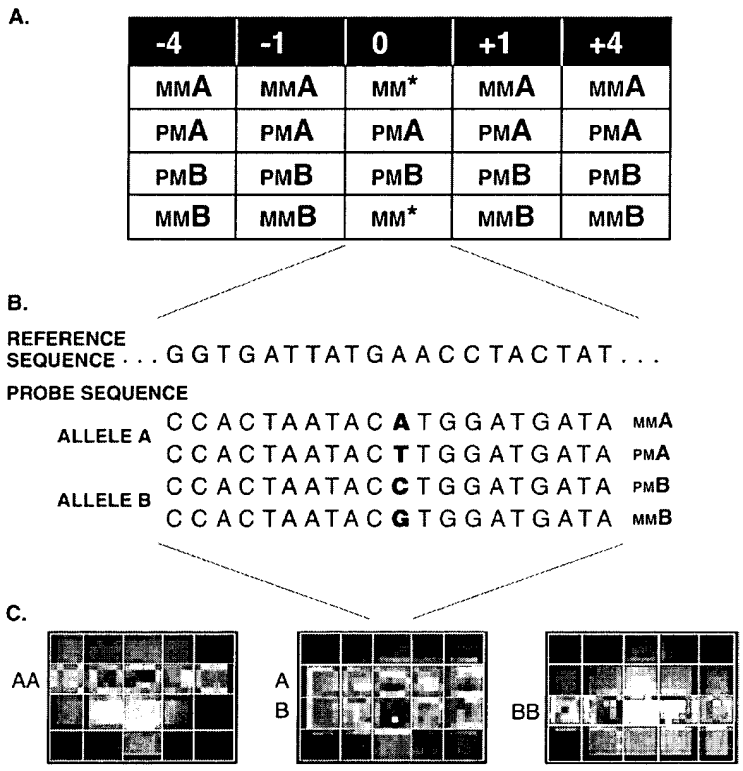


Fig. 12. HuSNP array design. **A** A known biallelic polymorphism at position 0 is interrogated by a block of four or five probe sets (five in this example). Each probe set consists of four probes, a perfect match and a mismatch to allele A, and a perfect match and a mismatch to allele B. One probe set in a block is centered directly over the polymorphism "0"), and others are centered upstream (−4, −1) and downstream (+1, +4). **B** The sequences of the probe set centered over the polymorphism is shown. **C** Sample images of blocks showing homozygous *A*, heterozygous *A/B*, or homozygous *B* at the same SNP site. (Reprinted with permission from Warrington JA et al (2000) In: Microarray biochip technology. Biotechniques Books, p 122)

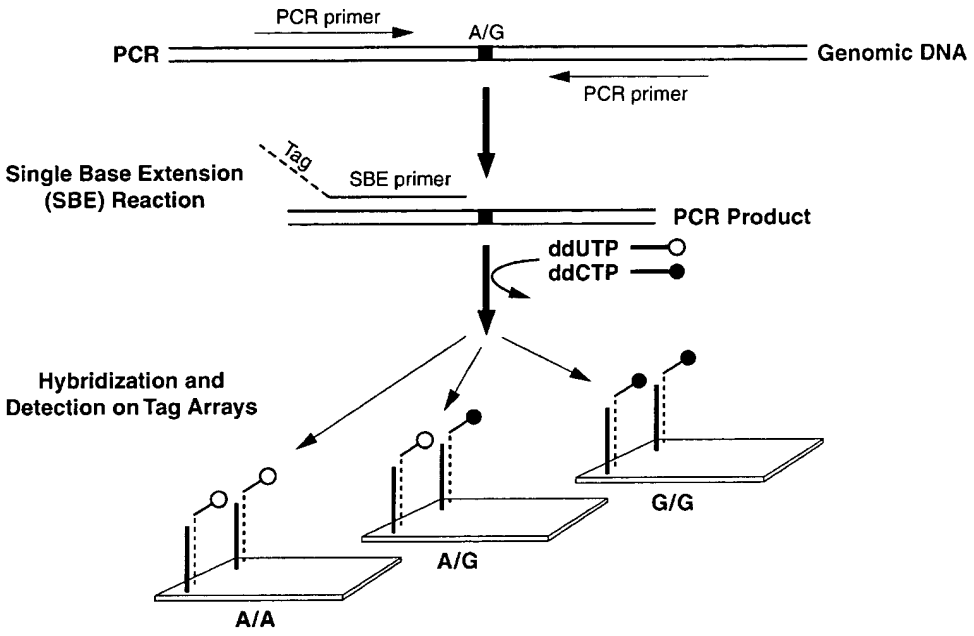


Fig. 13. Schematic of the single-base extension assay applied to Tag probe arrays. Regions containing known SNP sites (A or G in this example) are first amplified by PCR. The PCR product serves as the template for an extension reaction from a chimeric primer consisting of a 5′ tag sequence and a 3′ sequence that abuts the polymorphic site. The two dideoxy-NTPs that could be incorporated are labeled with different flurophors; in this example, ddUTP is incorporated in the case of the A allele, and ddCTP for the G allele. Multiple SBE reactions can be done in a single tube. The tag sequence, unique for each SNP, directs the extension products to a particular address on the Tag probe array. The proportion of a fluorophor at an address reflects the abundance of the corresponding allele in the original DNA. (Reprinted with permission from [45])

Although it is anticipated that the HuSNP assay will be appropriate for many applications, a more generic alternative is available in the form of the GenFlex array. For this array, 2000 20mer "tag" probe sequences were selected on the basis of uniform hybridization properties and sequence specificity. The array includes three control probes for each tag (a complementary probe and single-base mismatch probes for both the tag and its complement). One way to use the GenFlex array for SNP analysis is illustrated below (Fig. 13). In this example, a single-base extension reaction is used, in which a primer abutting the SNP is extended by one base in the presence of the two possible dideoxy-NTPs, each of which is labeled with a different fluorophor. Since each target-specific primer carries a different tag, the identity of each SNP is determined by hybridization of the single-base extension product to the corresponding tag probe in the GenFlex array [46]. The flexibility of the GenFlex approach lies in the freedom to partner any primer with any tag, a feature that enables other applications as well (Sect. 3.3).

Unlike yeast or bacteria, the size and complexity of the human genome currently necessitates locus-specific amplification for these genotyping applications. Without amplification, the concentration of each target is simply too low. We are developing more general arrays and procedures to reduce sequence complexity

while maintaining sufficient information content. This will eventually reduce or perhaps even eliminate the need to design, make, and handle large numbers of locus-specific primers and PCR products.

3.3 Other Applications

While oligonucleotide arrays have been used primarily to determine the composition of RNA or DNA, many other applications are possible as well. Any methodology that involves capturing large numbers of molecules that will hybridize to oligonucleotides can conceivably benefit from the highly parallel nature of these microarrays. Furthermore, the hybridized molecules, which are essentially libraries, can serve as a platform for subsequent analyses based on biochemical reactions. We describe below several recent "non-traditional" uses of GeneChip arrays, and suggest a number of other potential applications as well.

Tag arrays, such as the GenFlex array mentioned in the preceding section, have been used as "molecular bar-code" detectors [47–49]. In these experiments, mixtures of multiple yeast strains – each carrying a unique tag in its genome and having a different gene deleted – were subjected to a test such as drug treatment or growth in minimal medium, and then tag probe arrays were used to determine the proportion of each strain in the surviving population. As in gene expression and genotyping applications, the molecular bar-coding strategy takes advantage of the ability of probe arrays to selectively bind many different sequences in a complex mixture simultaneously. Parallel processing is not only faster and easier – it also minimizes the effect of variations in sample handling, thereby increasing the accuracy and precision of the measurements.

It is also envisioned that tag probe arrays will be useful for proteomics and other protein screening applications. For example, by attaching a different oligonucleotide sequence tag to each member of a group of proteins to be analyzed, hybridization would allow them to be arrayed in discrete locations on a chip for parallel screening. Proteins of interest would be identified by their position on the array. In one possible approach (Fig. 14), the tag is attached to the protein genetically by linking the tag to the mRNA and then translating the protein in such a manner that the protein remains associated with the mRNA, as is done in ribosome display to create and capture high affinity antibodies [50]. The protein-mRNA-tag complex is hybridized to the tag probe array, and screened for protein activity on the array. It is also conceivable that the proteins could be translated on the array, after hybridization. Genes of interest are recovered, either directly from the array or from another aliquot of the mRNA library, by PCR using the tag sequence for one primer and a common 3′ end sequence as the other primer. One use for such a system would be in directed evolution projects in which large gene libraries are made by cloning into cells, usually bacteria or yeast, followed by propagating and screening each clone individually for production of a protein with new or improved properties. The tag system would not only eliminate the need to transform and handle individual clones, but would also allow highly parallel screening because thousands of variants could be assayed simul-

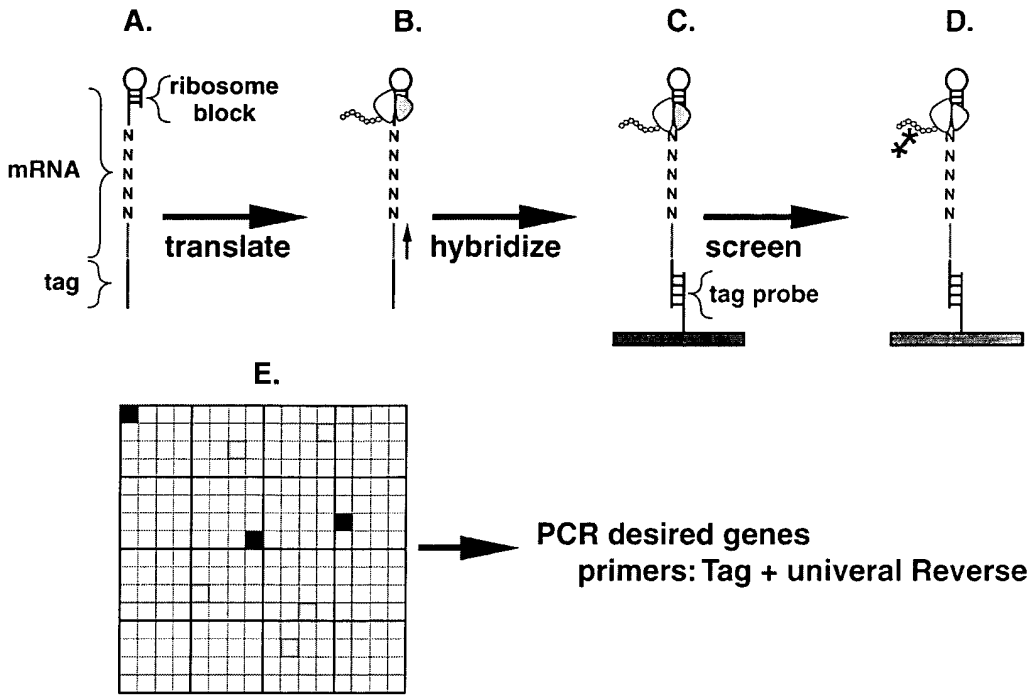


Fig. 14. Using Tag probe arrays to screen protein activity. To a protein-encoding mRNA a 5′ tag sequence and a 3′ ribosome-blocking sequence are attached (**A**). In a pool of such molecules, such as a randomly mutated gene library, each mRNA is paired with a unique tag and all have the same 3′ sequence. Following in-vitro translation either on a microarray or in a test tube, the nascent protein remains attached to the mRNA (**B**). During hybridization the tag directs each mRNA-protein to a particular address on the Tag probe array (**C**), where all the proteins are screened simultaneously for activity (**D**). Appropriate detection methods identify proteins of interest (**E**, *black and/or shaded blocks*). Finally, the corresponding genes can be captured by PCR of the mRNA pool using a universal reverse primer and each identified Tag sequence as a forward primer

taneously on the same array. Another use for the tag system would be to screen (poly)peptides made from existing mRNA molecules for properties such as drug binding. For example, all the mRNAs from a pathogenic bacterial strain could be converted to tagged proteins, which could then be screened for the ability to bind antibiotic candidates. The RNA molecules themselves could also be screened, as some drugs act directly on RNA. It is also conceivable that the oligonucleotide tag could be added directly to proteins, a method that might be useful in cases in which clones are already separated and one wishes to use the tag probe array only for parallel screening.

Researchers have also found a variety of novel uses for GeneChip arrays that were originally designed for gene expression monitoring. For example, Cho and co-workers [51] carried out a yeast two-hybrid study of *S. cerevisiae* proteins, mixing DNA from positive clones and hybridizing them to a yeast expression array, enabling them to identify the clones much more efficiently than if they had sequenced the clones by traditional methods. Winzeler et al. [52] used yeast expression arrays to identify more than 3700 biallelic genomic variations between

two yeast strains and then used the markers to simultaneously map five loci with high resolution (11–64 kb). Deletions in yeast [53] and in a clinical *M. tuberculosis* strain [54] were identified by similar techniques, a potentially important application given the propensity of some pathogens to avoid drug and vaccination treatments by deleting segments of their genomes. The use of multiple probes for each gene, a characteristic feature of GeneChip expression arrays, was critical to the high degree of resolution that was achieved in these experiments.

Milner et al. [55] reported using oligonucleotide arrays to survey oligonucleotide binding to a specific mRNA. A prevalent approach in anti-gene therapeutics involves knocking out malfunctioning genes through RNase H-mediated degradation of the mRNA, induced by duplex formation with antisense oligonucleotides. Presumably, the ability to predict, or at least empirically select, oligonucleotides which hybridize best to a given mRNA would be useful in the development of anti-gene therapeutics.

Gene expression arrays have also served as platforms for the analysis of splice variation in organisms with introns, and for mapping transcriptional boundaries [56]. Also, samples can be preselected for certain properties before hybridization, and at least two examples of this have been reported. In one case, cellular RNA [57], and in the other cellular DNA [58], were mixed with specific proteins, and complexes were purified by immunoprecipitation. Hybridization of the nucleic acids from the purified complexes revealed specific associations with the proteins. These two elegant experiments were carried out using arrays of spotted PCR products, but, again, one would expect that data of even higher resolution would be achievable using the multi-sequence probe sets present on GeneChip expression monitoring arrays.

One could conceivably treat hybridized expression arrays with RNase H, and assay for activity directly by following the loss of signal. This type of approach could be useful for revealing potential RNase H "hotspots" in mRNAs. A number of other powerful, but as yet under-utilized, applications also use arrayed probe-target duplexes as a platform for further reactions or analyses [59]. For example, Bulyk, Church and co-workers [60] created arrays of four base-pair restriction enzyme recognition sites and demonstrated activity with the appropriate enzymes, including *dam* methylase. Studies such as this provide further demonstration that arrays of double-stranded probes offer a promising platform for studying DNA-protein interactions. Methods based on template-directed enzymatic probe extension, ligation or cleavage are also being investigated as a means of improving allelic discrimination in hybridization-based mutation and polymorphism detection on arrays [16]. It is expected that hybridization-based biochemical assays on DNA microarrays will become increasingly commonplace in the coming years, especially in the area of high-throughput genotyping applications.

4 References

1. Phimister B (ed) (1999) Nat Genet Suppl 21:1
2. Schena R, Davis RW (2000) In: Schena M (ed) Microarray biochip technology. BioTechniques Books, Natick, MA, p 1
3. Fodor SPA, Read JL, Pirrung MC, Stryer LT, Lu A, Solas D (1991) Science 251:767
4. Pease AC, Solas D, Sullivan EJ, Cronin MT, Holmes CP, Fodor SPA (1994) Proc Natl Acad Sci USA 91:5022
5. McGall GH, Barone AD, Diggelmann M, Fodor SPA, Gentalen E, Ngo N (1997) J Am Chem Soc 119:5081
6. Lipshutz R, Fodor SPA, Gingeras TR, Lockhart DJ (1999) Nat Genet Suppl 21:20
7. McGall GH, Fidanza JA (2001) In: Rampal JB (ed) Methods in molecular biology. DNA arrays methods and protocols. Humana Press, Inc., Totowa, NJ, p 71
8. McGall GH, Barone AD, Diggelmann M (1999) Eur Pat Appl EP 967,217
9. Barone AD, Beecher JE, Bury P, Chen C, Doede T, Fidanza JA, McGall GH (2001) Nucleosides Nucleotides 20:525
10. Unpublished data
11. McGall GH (1997) In: Hori W (ed) Biochip arrays. IBC Library Series, Southboro, MA, p 2.1
12. McGall GH, Nam NQ, Rava R (2000) US Patent 6,147,205
13. Hasan A, Stengele K-P, Giegrich H, Cornwell P, Isham KR, Sachleben R, Pfleiderer W, Foote RS (1997) Tetrahedron 53:4247
14. Pirrung MC, Fallon L, McGall G (1998) J Org Chem 63:241
15. Beier M, Hoheisel JD (2000) Nucleic Acids Res 28:e11
16. Tonisson N, Kurg A, Lohmussaar E, Metspalu A (2000) In: Schena M (ed) Microarray biochip technology. BioTechniques Books, Natick, MA, p 247
17. McGall G, Labadie J, Brock P, Wallraff G, Nguyen T, Hinsberg W (1996) Proc Natl Acad Sci USA 93:13555
18. Wallraff G, Labadie J, Brock P, Nguyen T, Huynh T, Hinsberg W, McGall G (1997) Chemtech, February:22
19. Beecher JE, McGall GH, Goldberg MJ (1997) Preprints Am Chem Soc Div Polym Mater Sci Eng 76:597
20. Beecher JE, McGall GH, Goldberg MJ (1997) Preprints Am Chem Soc Div Polym Mater Sci Eng 77:394
21. Singh-Gasson S, Green RD, Yue Y, Nelson C, Blattner F, Sussman MR, Cerrina F (1999) Nat Biotechnol 17:974
22. Cho RJ, Campbell MJ, Winzeler EA, Steinmetz L, Conway A, Wodicka L, Wolfsberg TG, Gabrielian AE, Landsman D, Lockhart DJ, Davis RW (1998) Mol Cell 2:65
23. Cho RJ, Huang M, Dong H, Steinmetz L, Sapinoso L, Hampton G, Elledge SJ, Davis RW, Lockhart DJ, Campbell MJ (2001) Nat Genet 27:48
24. Debouck C, Goodfellow PN (1999) Nat Genet Suppl 21:48
25. Liotta L, Petricoin E (2000) Nat Rev Genet 1:48
26. Chen YW, Zhao P, Borup R, Hoffman EP (2000) J Cell Biol 151:1321
27. Wilson SB, Kent SC, Horton HF, Hill AA, Bollyky PL, Hafler DA, Strominger JL, Byrne MC (2000) Proc Natl Acad Sci USA 97:7411
28. Lee CK, Klopp RG, Weindruch R, Prolla TA (1999) Science 285:1390
29. Ly DH, Lockhart DJ, Lerner RA, Schultz PG (2000) Science 287:2486
30. Alon U, Barkai N, Notterman DA, Gish K, Ybarra S, Mack D, Levine AJ (1999) Proc Natl Acad Sci USA 96:6745
31. Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, Coller H, Loh ML, Downing JR, Caligiuri MA, Bloomfield CD, Lander ES (1999) Science 286:531
32. Chee M, Yang R, Hubbell E, Berno A, Huang XC, Stern D, Winkler J, Lockhart DJ, Morris MS, Fodor SP (1996) Science 274:610

33. Hacia JG, Sun B, Hunt N, Edgemon K, Mosbrook D, Robbins C, Fodor SP, Tagle DA, Collins FS (1998) Genome Res 8:1245
34. Hacia JG, Brody LC, Chee MS, Fodor SP, Collins FS (1996) Nat Genet 14:441
35. Hacia JG (1999) Nat Genet 21:42
36. Ahrendt SA, Halachmi S, Chow JT, Wu L, Halachmi N, Yang SC, Wehage S, Jen J, Sidransky D (1999) Proc Natl Acad Sci USA 96:7382
37. Wikman FP, Lu ML, Thykjaer T, Olesen SH, Andersen LD, Cordon-Cardo C, Orntoft TF (2000) Clin Chem 46:1555
38. Liu WW, Webster T, Aggarwal A, Pho M, Cronin M, Ryder T (1997) Am J Hum Genet 61:1494
39. Kozal MJ, Shah N, Shen N, Yang R, Fucini R, Merigan TC, Richman DD, Morris D, Hubbell E, Chee M, Gingeras TR (1996) Nat Med 2:753
40. Gunthard HF, Wong JK, Ignacio CC, Havlir DV, Richman DD (1998) AIDS Res Hum Retroviruses 14:869
41. Wang DG, Fan JB, Siao CJ, Berno A, Young P, Sapolsky R, Ghandour G, Perkins N, Winchester E, Spencer J, Kruglyak L, Stein L, Hsie L, Topaloglou T, Hubbell E, Robinson E, Mittmann M, Morris MS, Shen N, Kilburn D, Rioux J, Nusbaum C, Rozen S, Hudson TJ, Lander ES, et al (1998) Science 280:1077
42. Cargill M, Altshuler D, Ireland J, Sklar P, Ardlie K, Patil N, Shaw N, Lane CR, Lim EP, Kalyanaraman N, Nemesh J, Ziaugra L, Friedland L, Rolfe A, Warrington J, Lipshutz R, Daley GQ, Lander ES (1999) Nat Genet 22:231
43. Lindblad-Toh K, Winchester E, Daly MJ, Wang DG, Hirschhorn JN, Laviolette JP, Ardlie K, Reich DE, Robinson E, Sklar P, Shah N, Thomas D, Fan JB, Gingeras T, Warrington J, Patil N, Hudson TJ, Lander ES (2000) Nat Genet 24:381
44. Lindblad-Toh K, Tanenbaum DM, Daly MJ, Winchester E, Lui WO, Villapakkam A, Stanton SE, Larsson C, Hudson TJ, Johnson BE, Lander ES, Meyerson M (2000) Nat Biotechnol 18:1001
45. Mei R, Galipeau PC, Prass C, Berno A, Ghandour G, Patil N, Wolff RK, Chee MS, Reid BJ, Lockhart DJ (2000) Genome Res 10:1126
46. Fan JB, Chen X, Halushka MK, Berno A, Huang X, Ryder T, Lipshutz RJ, Lockhart DJ, Chakravarti A (2000) Genome Res 10:853
47. Shoemaker DD, Lashkari DA, Morris D, Mittmann M, Davis RW (1996) Nat Genet 14:450
48. Giaever G, Shoemaker DD, Jones TW, Liang H, Winzeler EA, Astromoff A, Davis RW (1999) Nat Genet 21:278
49. Winzeler EA, Shoemaker DD, Astromoff A, Liang H, Anderson K, Andre B, Bangham R, Benito R, Boeke JD, Bussey H, Chu AM, Connelly C, Davis K, Dietrich F, Dow SW, El Bakkoury M, Foury F, Friend SH, Gentalen E, Giaever G, Hegemann JH, Jones T, Laub M, Liao H, Davis RW, et al (1999) Science 285:901
50. Hanes J, Jermutus L, Pluckthun A (2000) Methods Enzymol 328:404
51. Cho RJ, Fromont-Racine M, Wodicka L, Feierbach B, Stearns T, Legrain P, Lockhart DJ, Davis RW (1998) Proc Natl Acad Sci USA 95:3752
52. Winzeler EA, Richards DR, Conway AR, Goldstein AL, Kalman S, McCullough MJ, McCusker JH, Stevens DA, Wodicka L, Lockhart DJ, Davis RW (1998) Science 281:1194
53. Winzeler EA, Lee B, McCusker JH, Davis RW (1999) Parasitology 118:S73
54. Salamon H, Kato-Maeda M, Small PM, Drenkow J, Gingeras TR (2000) Genome Res 10:2044
55. Milner N, Mir KU, Southern EM (1997) Nat Biotechnol 15:537
56. Wassarman KM, Repoila F, Rosenow C, Storz G, Gottesman S (2001) Genes Dev 15:1637
57. Takizawa PA, DeRisi JL, Wilhelm JE, Vale RD (2000) Science 290:341
58. Ren B, Robert F, Wyrick JJ, Aparicio O, Jennings EG, Simon I, Zeitlinger J, Schreiber J, Hannett N, Kanin E, Volkert TL, Wilson CJ, Bell SP, Young RA (2000) Science 290:2306
59. Gunderson KL, Huang XC, Morris MS, Lipshutz RJ, Lockhart DJ, Chee MS (1998) Genome Res 8:1142
60. Bulyk ML, Gentalen E, Lockhart DJ, Church GM (1999) Nat Biotechnol 17:573

Received: June 2001